

<https://helda.helsinki.fi>

Aspects of the Institutionalization of Evaluation in Finland : Basic, Agency, Process and Change

Ahonen, Pertti

2015-07-07

Ahonen , P 2015 , ' Aspects of the Institutionalization of Evaluation in Finland : Basic,
Agency, Process and Change ' , Evaluation , vol. 21 , no. 3 , pp. 308-324 . <https://doi.org/10.1177/1356389015592546>

<http://hdl.handle.net/10138/232564>

<https://doi.org/10.1177/1356389015592546>

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

**Evaluation in its basic, agency, process and change aspects: General considerations
and a country case study on the institutionalization of evaluation**

Pertti Ahonen

University of Helsinki, Finland

Pertti Ahonen (DSocSc, MScEcon) is professor of political science at the University of Helsinki, and an evaluation scholar and practitioner since thirty-five years.

Corresponding author:

Pertti Ahonen, Department of Political and Economic Studies, Unioninkatu 37, 00014

University of Helsinki, Finland

Email: pertti.ahonen@helsinki.fi

Evaluation in its basic, agency, process and change aspects: General considerations and a country case study on the institutionalization of evaluation

Abstract

This article utilizes selected neo-institutional theoretical ideas to approach evaluation. Empirically, the article examines the institutionalization of evaluation in the national government of Finland. The results indicate ambiguity in the basic institutionalization of the Finnish evaluation, and imprecision in the agency of the actors that carry out or commission evaluations or utilize the evaluation results. Some of the Finnish institutional practices of evaluation enhance formal rationality such as efficiency and effectiveness, some support legitimation, and others do both in combination, and the strength of the coupling of evaluation to decision-making varies greatly. For future research, the article suggests studies on the institutionalization of evaluation in other countries. For evaluation practice, the results emphasize attention to the position of evaluation in the rationality-legitimation axis, and to the variable coupling of evaluation to decision-making.

Keywords

Institutionalization, agency, rationality, efficiency, effectiveness, legitimation, coupling,
institutional change, Finland

Evaluation comprises a practical art but one that widely utilizes varieties of science and often contributes substantially to these varieties. Because scientific approaches to evaluation are many and different, this article drops pretensions to open more than one perspective upon evaluation. The article derives its conceptual framework from global neo-institutional research (Peters, 2011; Lowndes and Roberts, 2013). On theoretical grounds that the next section explicates, the article poses four research questions and empirically seeks answers to these questions in a country case study on the institutionalization of evaluation in Finland:

- (1) Assuming that evaluation has deeper institutional roots than first appear, what do the processes and outcomes of its basic institutionalization comprise of?
- (2) Assuming that the activities of actors engaged in evaluation have deeper roots than first appear, from where do these actors derive their capacity to act – their ‘agency’?
- (3) Observing the global proliferation and the resilience of evaluation, which processes of institutionalization contribute to its shaping and persistence?
- (4) Observing that evaluation may not only evolve in longer-term processes in a vegetative way, how do active change agents influence its shaping?

In principle, academic research on evaluation and evaluation practice may manage without reflecting upon the key notions of evaluation, taking these notions for granted. Guba and Lincoln (1989), in their turn, advise us to abandon the efforts to define evaluation on the expectancy that context-free definitions turn vacuous. However, to have our research texts on evaluation published in the first place we may have to offer universal or contextual definitions of evaluation, and in our capacity as practical evaluators we may have to fulfill analogous requests of our customers (see, for instance, Stufflebeam and Shinkfield, 2007). The entrenched capacity of a given variety of neo-institutional research to examine the fluid dynamics of meanings and determinations has recommended this variety for this study.

The following section elaborates the conceptual approach, gives motivation for the four research questions, and links these questions to each other. The next section explicates the research methodology. Each of the subsequent four sections seeks answers to one of the questions. The last section discusses the contributions and implications of the study.

The research approach

Within neo-institutional research (Peters, 2011; Lowndes and Roberts, 2013), this article represents what is commonly called sociological neo-institutionalism in a

subspecies launched by John W. Meyer and Brian Rowan (1977) during the incubation years of any neo-institutionalism to speak of, and later substantially elaborated by Meyer and his colleagues (Kruecken and Drori, 2009). This article motivates its four research questions by means of four complementary perspectives it draws out from the Meyerian neo-institutionalism. The first perspective grounds the examination of the roots and processes of basic institutionalization, the second one deepens on ‘agency’ as the capacity of the institutional actors to act, the third one focuses upon broader processes and outcomes of the ‘going concern’ of institutionalization, and the last perspective concerns intense agent-driven institutional change.

The basic institutionalization of evaluation

The neo-institutionalism that this article draws upon sees basic institutionalization to be comprised of the formation of habitually taken-for-granted patterns of interaction between the relevant actors, catalyzing also the valid social knowledge that these actors share (Meyer and Rowan, 1977: 341). Zucker (1977: 726) reminded during the incubation years of neo-institutionalism that this ‘social knowledge, once institutionalized, exists as a fact, as part of objective reality’. The neo-institutionalism in hand seeks such facts in sources including institutional vocabularies, terminologies,

conceptual systems, classifications, categorizations, boundaries within and between institutions, and the identities of institutions and the institutional actors.

Institutional habituation poses particular challenges to the methodology of research, as it requires caution in using institutional insiders as informants. Schneiberg and Clemens (2006: 214) suggest that we examine ‘breaches’, ‘deviant events, or conflicts that reveal ... undiscussed boundaries of taken-for-granted understanding’. The intricacies of institutional habituation attract this article to dive below the surface of evaluation to examine its ambiguity and its hybrid forms, its equivocal language, its leaking classifications, categorizations and boundaries, and the incomplete identities of evaluators, evaluation practices, and institutions of evaluation. Emphasizing these characteristics by no means alleges that evaluation has fundamental faults, but indeed only comprises a methodological point.

The capacity to act in evaluation: Evaluation agency

This article will advance from the basic institutionalization of evaluation to examine the ‘evaluation agency’ of actors in real-world contingent circumstances in which these actors may either to succeed or fail. Meyer and Jepperson (2000: 117) analytically distinguish three types of agency. ‘Agency for itself’, in the plural ‘agency for themselves’, is vested in individuals and collectivities such as professions, organizations

and states. ‘Agency for others’ is carried out, for instance, by professions, specialized organizations and nation states on behalf of individuals, families, groups, organizations and other states. ‘Agency for standards and principles’ represents, for example, human rights, transparency, good governance, social responsibility, and science-driven rationalization (Drori and Meyer, 2006; Drori et al., 2009).

This article looks for the first type of agency in the self-evaluation of actors and the evaluation of evaluatees by evaluators. The article looks for agency for others in evaluations that evaluators carry out on commission from other actors, and in evaluations carried out by professions, organizations and institutions of evaluation following their mission. Finally, the article traces agency for standards and principles in the very approaches, practices and principles of evaluation itself.

Diffusion, rationality, legitimation, and institutional coupling in evaluation

The neo-institutionalism that this article utilizes keenly examines the diffusion of models and scripts for new or revised institutional elements, the modification of these models and scripts in their uses, and the sedimentation of the new elements atop those that have been institutionalized earlier (Strang and Macy, 2001). The article sees evaluation to comprise one of the phenomena to examine in these very terms.

Accounts on institutional actors trying to control their contingent circumstances by means of institutional elements that enhance formal rationality such as efficiency, performance or effectiveness convey only a partial truth (Meyer and Rowan, 1977). This is because the actors may also introduce and maintain institutional elements that enhance institutional legitimation.

Neo-institutionalism agrees with other orientations of research that if institutional elements introduced to enhance rationality fulfill their purpose, this supports institutional legitimation. However, since Meyer and Rowan (1977), neo-institutionalism examines also the possibility that legitimation receives enhancement from institutional elements introduced in the name of *promising* rationality independently of its *actual* achievement. This does not necessarily indicate a deficiency, but decoupling or loose coupling between institutional elements that actually enhance rationality and those that enhance legitimation may usefully deny institutional outsiders from access to the vulnerable rational institutional core. Research on evaluation recognizes the complex relationships between rationality and legitimation years before neo-institutional research arose (Eaton, 1962; Weiss 1970; but see also Dirsmith et al., 2000).

Intense agent-driven institutional change in evaluation

Lowndes and Roberts (2013: 111–43) distinguish two dimensions of institutional change, comprised of its tempo and the strength of the agency of the actors that initiate change. Combining these two dimensions, they distinguish four types of change. The other types have been covered in the previous section by and large, but the fourth type is comprised of intense agent-driven institutional change (Lowndes and Roberts, 2013: 122–5). In examining the institutionalization of evaluation, this article subscribes to the early neo-institutional view that Zucker (1977: 726) elaborated:

(W)hen emergent culture is the focus, then the problem of establishing facticity becomes the central problem. It is here that the moral character of social facts becomes the central concern.

In the neo-institutional understanding to which this article subscribes, the actors that successfully carry agency for change induce the re-evaluation of what count for the best acceptable social facts as concerns the institutionalization of evaluation. This re-evaluation substitutes a new preferred institutionalization for the predecessor institutionalization that turns unacceptable (see, for instance, Meyer and Hammerschmid, 2006).

Methodology

Common criteria of qualitative research oblige this article: the cogent formulation and theoretical motivation of the research questions, the sufficient explication of the material collection and analysis, the reasonable saturation of the material in the analysis, and, for the readers to assess, the credibility and trustworthiness of the results (Silverman, 2011: 27–56). Empirically, the article examines the institutionalization of evaluation within Finland’s national government of twelve ministries and numerous agencies and offices (Salminen et al., 2012).

Had Finland a Westminster-style government with a more unitary political executive, this article should amply draw upon such sources as white papers elaborating government policies and programs. However, Finland’s legalist hierarchy of norms, descending from a written constitution downwards, justifies reliance on the contents of the national legislation database *Finlex*. The article will first use this database to examine the basic institutionalization of evaluation. Abandoning essentialist search for evaluation under linguistic expressions that shy away from this notion, the article advances in a nominalist mode, looking for instances of the Finnish word *arviointi* for ‘evaluation’, and searching for ambiguity in the use of this word.

The material to examine institutional agency for evaluation in Finland includes official documents and research studies. Analogous material and *Finlex* support the examination of diffusion, rationality versus legitimation and coupling in the

institutionalization of evaluation. The study of intense agent-driven institutional change in evaluation requires, in principle, the observance of methodological conventions of historical research (Jordanova, 2006). However, the article genre allows no more than examples of cases with such change.

Besides documents, the research material includes five expert interviews to check on facts and tease out interesting ‘breaches’ in the institutionalization of evaluation. Following certain conventions of qualitative interviewing, the interview questions were not shown the interviewees but only used as prompts in each interview situation (Gubrium et al., 2012; Appendix 2). All interviewees had experience of evaluation-related research and evaluation practice, each held a doctorate or was earning it, and the survey included both males and females. The mature average age of the interviewees was fifty-five. To encourage the interviewees to speak, the interviews were carried under a binding promise of anonymity.

Basic institutionalization of evaluation in Finland

In the Finnish legislative database *Finlex* as in the Finnish language in general, the word *arviointi* somewhat unfortunately may signify not only ‘evaluation’ but also, for instance, ‘assessment’, ‘inspection’, ‘valuation’, ‘ascertainment’, or mere ‘checking’ of legal compliance. Nominally, *arviointi* is present in quite a number of acts and statutes

(Table 1), but deriving from the polysemy, most instances have little to do with evaluation as understood within the global evaluation community. The common Finnish expression *tuloksellisuuden arviointi*, despite its literal translation as ‘performance evaluation’, refers to performance *measurement* ascribing quantitative values to predefined performance indicators. Moreover, the majority of the current 315 acts or statutes that use the expression *vaikuttavuuden arviointi* (‘effectiveness evaluation’; see Table 1) comprise of nothing but upgrades of technical ‘inspection’, ‘environmental impact assessment’, or other impact assessment. Some varieties of assessment are closer than second cousins to evaluation, but methodological nominalism requires indicating differences between the two notions.

Table 1

The core of the Finnish legislative vocabulary of effectiveness evaluation proper can be summed up in the notion *yhteiskunnallisen vaikuttavuuden arviointi*. Its English translation, the ‘evaluation of societal effectiveness’, is no less awkward than the Finnish original. The notion was found in 42 current acts or statutes (Table 1). However, some of the 42 do not really deal with ‘societal effectiveness’, and several others regulate various types of impact assessment most of which lack undisputable elements of evaluation. Eliminations left three solid domains to examine closer. One

includes the universities, removed from the government budget and the civil service in 2010, but with a statutory obligation to pursue ‘societal effectiveness’. The working order of the state council, comprised of the ministers and the twelve ministries together, also prescribes the evaluation of societal effectiveness, as does the working order of each ministry. The third domain was established in early 2000s legislation (BA, 2013; BS, 2013), instituting the evaluation of societal effectiveness in the annual budget preparation, the medium-range planning in each ministry for its sector of administration, the annual reporting of the agencies and offices, the annual reporting of the ministries on their sectors, and the annual government report to Parliament on the government accounting closure.

During my research process colleagues continued asking why not use other sources except for *Finlex* to examine the uses of the Finnish word for ‘evaluation’, *arviointi*. Besides Finland’s entrenched legalist traditions, the results of an earlier study (Ahonen and Virtanen, 2005) motivate the chosen concentration. By middle of the first decade of the 2000s, *arviointi* had received very wide interpretations in this country, as many earlier types of study, research or monitoring were renamed *arviointi* but without introducing changes in the substance. Had this article exploited other documents except for legal texts, it might have reaped chaos. Future studies find here their change, such as studies using big data methods to dredge the hundreds of the annual Finnish evaluation reports.

Agency for evaluation in Finland

Agency of the evaluation actors for themselves

Let us not allow ambiguities in the basic institutionalization of evaluation in Finland to discourage us from examining agency for evaluation in Finland. We cannot find a master evaluator institution in this country. The State Audit Office has the independence, but lacks the statutory mandate to rename its performance auditing, often paralleling evaluation, to comprise evaluation proper (Appendix 1, item 1.2). Despite efforts, no assertive coordination of evaluation has arisen within Finland's national government (Appendix 1, item 2; PMO, 2011; MF, 2012).

Appendix 1 implies rather than explicitly reveals that the characteristic evaluators within Finland's national government comprise employees of agencies and offices subordinated to the ministries rather than officials in the ministries themselves. Many other evaluators work for global and domestic consulting companies and other organizations that win evaluation commissions from government organizations.

Evaluation agency on behalf of others

The constitutional obligation of Finland's government to respond to parliamentary requests for comprehensive ex post reporting on the results of legislative reforms establishes an important type of the evaluation agency in this country (Appendix 1, item 1). Further types of evaluation agency are written in legislation on individual policy fields, result contracts between the ministries and their subordinate agencies and offices, and business contracts between outsourced providers of evaluation services and government organizations.

My interviewees held the opinion that no profession proper has evolved in Finland to carry general agency for evaluation. Since 1992, those who intend to acquire Finland's certified degree of auditors of public administration and the public finances (JHTT) have to pass evaluation assignments in their examination, but most of the JHTTs play few evaluation roles. Since its establishment in 1999, the Finnish Evaluation Society (FES), with approximately 200 members, has made valuable contributions. However, it has not forged a unitary body of evaluators, and many evaluators have never joined the society.

At the time of this study, no other than micro or small companies indicated exclusive specialization in evaluation in Finland, but reflecting the small evaluation market, many larger companies offered evaluation expertise alongside of other services. However, the Finnish markets have not been too small to attract foreign companies to acquire evaluation operations first evolved under Finnish ownership.

Evaluation agency representing general principles and standards

Neither the members of Finland's evaluation community nor the government have passed evaluation standards and principles of their own. One of the interviewees argued that competitive evaluation markets can be defined between the approximately one hundred organizations commissioning evaluations in Finland's national government on the one hand, and on the other the providers of evaluation services. The interviewee proposed also that the market competition provides better quality assurance than formal standards and principles.

Finland's evaluators frequently apply global standards and principles (for instance, AEA, 2013), and some government organizations have explicitly harmonized their evaluation guidelines with these standards and principles. As the earliest case, for several decades already the evaluation of Finland's development co-operation (Appendix 1, item 3) subscribes to global and since the 1990s also to EU evaluation standards and principles (MFA, 2013).

Softer and more academic standards and principles that Finland's scholars and practitioners of evaluation apply derive from all three types of 'evaluation roots' that Alkin (2012) distinguishes. Evaluation sprouting from the 'methods' roots evolved from early quasi-experiments on speed limits (Salusjärvi and Kontiala, 1975) in line with the

evaluation classic of Campbell and Ross (1968). Evaluative studies of school achievement also sprang up early on (Renko, 1971). Certain non-experimental and experimental economic analyses have been innovatively re-named evaluations (Ilmakunnas et al., 2008), and evidence-based evaluation within health care applies entrenched conventions of randomized control trials (Hakama et al., 2012). Some of Finland's education scholars have stressed Alkin's 'valuing' roots of evaluation (Keskitalo et al., 2012), and Alkin's 'utilization' roots (Patton, 2008) have also won representation (Berg and Hukkinen, 2011). Harder to classify evaluation types include 'fourth-generation evaluation' (Guba and Lincoln, 1989; García-Rosell and Mäkinen, 2013), evaluation that implements critical scientific realism (Pawson and Tilley, 1997; Holma and Kontinen, 2010), and studies in evaluation ethics (Schwandt, 2002; Laitinen, 2008).

Numerous doctoral dissertations have been written in Finland on evaluation since the earliest study (Sintonen, 1981), but evaluation is not among the country's approximately 300 academic disciplines awarding master's and doctoral degrees, nor the main theme in any of the thematic master's programs, which also number about 300. My interviewees vented their frustration over too common academic reluctance to assign full intellectual value to research representing evaluation. However, the vigorous global publication activity by Finland's evaluation scholars questions the claims of deficiency.

Diffusion, performance, legitimation and coupling in the institutionalization of the Finnish evaluation

After considering the basic institutionalization of evaluation and agency for evaluation in Finland, let us have a look at the ‘going concern’ of the Finnish evaluation in its practical institutionalization. As the conceptual framework implies, the examination has to provide indications of the diffusion, modification and sedimentation of global institutional elements of evaluation in the Finnish case (Appendix 1, last column).

Responding to a collegial suggestion, this article will also take a look at the enhancement of rationality and the enhancement of legitimation by means of evaluation, and at the application of institutional tight or loose coupling between evaluation and actual decision-making. Besides evaluation proper, the examination includes practices close enough to evaluation to deserve consideration.

High politics, the elevated authority of institutions and the high political importance of issues crowd out rationality and accentuate legitimation as concerns evaluation (Appendix 1, items 1 and 1.1). In one case, the relative shortness of the electoral cycle – four years at the maximum – appears to crowd out rationality (item 2.0.1), in another the proximity to high politics exerts an analogous impact (items 2.0.2 and 2.0.3), and in a third case, limited operationalization restricts rationality (item 4). Within higher education policy legitimation receives emphasis in evaluation (item 8.1.2). **Spillovers**

from the resilient excellent results in evaluating the efficiency and the effectiveness of compulsory education in Finland also provide this country's system of compulsory education general legitimation (item 8.1.1). Finland's foremost impact assessment practice builds on institutionalized statutory procedures enacted in the *name* of rationality enhancement, but does necessarily ensure *actual* rationality but rather guarantees normative compliance (item 13). However, rationality and legitimation commonly coexist as emphases of evaluation.

The coupling of evaluation to decision-making is generally loose rather than tight in Finland's national government. However, tight and loose coupling commonly co-exist (items 1, 1.3, 3, 7.0.1, 9, 10, 12 and 13). The coupling is looser in high politics (items 1, 2.0.1 and 2.0.2), if operationalization is limited (items 4 and 7.0.2), or if constitutional or other statutory separation between legislative and executive powers exerts an influence (items 1.2, 1.3, and 7.0.2).

Insofar as organizations or functions have a staff status as opposed to a line status, this weakens the coupling of their evaluation results to the decision-making of the government ministries (items 4.1, 7.1, 8.1.3, 8.2, 9, 11.1, 12, 12.1, 12.2 and 13.1). In one case, an official prohibition prevails to publish other than summary evaluation results explicitly to prevent their tight coupling to decision-making on individual schools (item 8.1.1). Vested interests of the corporatist variety may also influence the coupling of evaluation to decision-making (12.2). Last, let us note newer government

efforts to tighten the coupling of functions with evaluation components to government decision-making (items 2.0.3. and 2.0.4), including the introduction of a new assessment practice of tight rather than loose coupling (item 2.0.3).

Intense agent-driven institutional change in Finland's national government evaluation

Next, let us turn from the 'going concern' of institutionalized evaluation in the Finnish national government to examining intense agent-driven institutional within evaluation. Finland turned into a donor of development co-operation in the 1960s, which invited the very first institutionalization of evaluation proper in the country to legitimate allocations from the government budget for the co-operation purposes (Appendix 1, item 3). This article has referred above also to two types of evidence-based applied research *avant la lettre* paralleling evaluation, comprised of quasi-experimental studies on speed limits and studies on teacher efficiency (Salusjärvi and Kontiala, 1975; Renko, 1971).

In 1987–91, evaluation was by and large overshadowed in Finland by New Public Management emphases on operative performance, and in 1991–94, by the exigencies of preventing the general collapse of the economy (on this, see the chapter on Finland in Pollitt, 2013). The start of Finland's membership in the EU in 1995 catalyzed the ex ante evaluation of the structural fund programs and projects (Appendix, item 11). Later,

ex ante evaluation expanded in legal policy-making and legal preparation (items 4 and 4.1; Tala, 2010). Finland harmonized the evaluation of its development co-operation and its research and technological development policies with their EU counterparts by 1995 at the latest (items 3, 8.2 and 11.2; MFA, 2013). Evaluations required by the open method of coordination of EU policies, evolved since 1997, have been also implemented in Finland (Ahonen and Virtanen, 2008). Since 2015, Finland implements systematic national ex ante assessments of the impact of European Commission legal and policy initiatives in this country (Uusikylä et al., 2015; Appendix, item 2.0.3., and see also items 2.0.4–2.0.5).

An important threshold year for evaluation in Finland situates itself in 1995, preceding the multiplication in types and targets of evaluations and the numbers evaluation studies in ways that suggest partial substitution of post-NPM procedures for those of NPM proper (see Pollitt and Bouckaert, 2011). Evaluation expanded, for instance, to cover national public administration reforms (Pollitt et al., 1997), public services, individual government organizations, and policies, programs and projects. However, as indicated in the first empirical section, the word *arviointi* (‘evaluation’) became a buzzword connected with many pursuits many that had little to do with evaluation as understood within the global evaluation community.

In the future, focused studies should examine further cases with intense agent-driven change in the Finnish evaluation scene. This concerns, for instance, the 2003–4 reform

that introduced the statutory evaluation of societal effectiveness (MF, 2003; BA, 2014; BS, 2014). This evaluation evolved not only an extension of government accounting, but also a novel dimension of budgetary accountability, and a new content element in the annual government accounting closure. The moral of the reform comprised substituting an effectiveness orientation for the traditional compliance orientation, and continuing to supplement the efficiency orientation of NPM with further post-NPM characteristics.

Contributions and implications

This article has sought answers to four research questions derived from a certain subspecies of neo-institutional theory (receiving its best overview in Kruecken and Drori, 2009) in its examination of the institutionalization of evaluation in the national government of Finland. The article has contributed with a country case study on the institutionalization of evaluation grounded on a systematic theoretical framework. The empirical results indicate ambiguity in the basic institutionalization of evaluation in Finland and in the agency of the actors carrying out, commissioning or using evaluations. However, evaluation principles, standards, approaches and methods help give coherence. With tighter or looser coupling to decision-making, some of the Finnish institutional practices of evaluation enhance formal rationality such as efficiency and

effectiveness, some support legitimation, and the rest contribute to both ends. Intense agency-drive change has not been missing in the Finnish evaluation scene, either.

A referee usefully asked if evaluation in Finland's national government is institutionalized at all. One of my interviewees helps provide the answer, proposing that the evaluation is institutionalized up to strong *routinization*, but characteristically confined to 'particular knowledge practices' (*erityiset tietokäytännöt*). The referee comment invited minor 'process tracing' (Goertz and Mahoney, 2012), deriving particularities of the Finnish case from the broader institutionalization that the OECD (2010) calls 'sector stovepipes', each managed by one of the government ministries. In general, the neo-institutionalism that this article applies fully agrees with views that global pressures towards isomorphism as concerns institutional forms – including those of evaluation – may not lead to homogeneity because of the wide variety of contexts in which these forms find their applications (see Beckert, 2010; Pollitt, 2013).

At its beginning, this article promised to open up one and only one perspective from among the others to examine the institutionalization of evaluation. However, this limitation by no means rules out examinations on the institutionalization of evaluation in other countries than Finland in analogous theoretical and methodological terms as in this article.

The implications of this article for practice are possibly stronger and more general than its immediate implications for future research. The examination of the basic

institutionalization advises the evaluation practitioner to try to live with ambiguity both as concerns notions of evaluation and the agency of the evaluator, those who commission evaluations, and those who use evaluation results. A clear-headed evaluator should find good use for the rationality versus legitimation division in evaluation and the various degrees of coupling between evaluation and decision-making. Last but not least, would not many a researcher on evaluation and many an evaluation practitioner want to turn into an agent of intense change for the very best of evaluation?

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- AEA (American Evaluation Association) (2013) Program evaluation standards. URL: <http://www.eval.org/p/cm/ld/fid=103>
- Ahonen P and Virtanen P (2005) Unpublished book manuscript with an empirical study of evaluation in Finland, in Finnish. Tampere: University of Tampere.
- Ahonen, P and Virtanen P (2008) Understanding evaluation in the social sector: two reflexive metaevaluation case studies inspired by Pierre Bourdieu. *International Journal of Public Administration* 31(10–11), 1146–66.

- Alkin MC (ed.) (2012) *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. 2. ed. Thousand Oaks, CA: Sage.
- BA (Budget Act) (2014) Originally 1988, with later amendments. In Finnish. URL: <http://www.finlex.fi/fi/laki/ajantasa/1988/19880423>
- Beckert, Jens (2010) Institutional isomorphism revisited: convergence and divergence in institutional change. *Sociological Theory* 28(2), 150–166.
- Berg A and Hukkinen JI (2011) Beyond effectiveness: the uses of Finland's national programme to promote sustainable consumption and production. *Journal of Cleaner Production* 19(16), 1788–97.
- BS (Budget Statute) (2014) 1992, with later amendments. In Finnish. URL: <http://www.finlex.fi/fi/laki/ajantasa/1992/19921243>
- Campbell HT and Ross DL (1968) The Connecticut crackdown on speeding: time-series data in quasi-experimental analysis. *Law & Society Review* 3(1): 33–54.
- Dirsmith MW, Fogarty TJ and Gupta P (2000) Institutional pressures and symbolic displays in a GAO context. *Organization Studies* 21(3), 515–37.
- Drori GS and Meyer JW (2006) Scientization: making the world safe for organizing. In: Djelic M-L and Sahlin-Andersson K (eds) *Transnational Governance*. Cambridge: Cambridge University Press, 33–52.

- Drori, GS, Meyer JW and Hwang H (2009) Global organization: rationalization and actorhood as dominant scripts. In: Meyer RE, Sahlin K, Ventresca M and Walgenbach P (eds) *Ideology and Institutions*. Bingley: Emerald, 17–43.
- Eaton JW (1962) Symbolic and substantive evaluative research. *Administrative Science Quarterly* 6(4), 421–42.
- ENQA (European Association for Quality Assurance in Higher Education) (2013) Homepage. URL: <http://www.enqa.eu>
- Finlex (2012) Database of Finland's legislation. URL: <http://www.finlex.fi>
- García-Rosell J-C and Mäkinen J (2013) An integrative framework for sustainability evaluation in tourism: applying the framework to tourism product development in Finnish Lapland. *Journal of Sustainable Tourism* 21(3), 396–416.
- Goertz, G. and Mahoney G (2012) *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton, NJ: Princeton University Press.
- Guba EG and Lincoln YS (1989) *Fourth Generation Evaluation*. Newbury Park, CA: Sage.
- Gubrium, J.F., Holstein, J.A., Marvasti, A.B. & McKinney, K.D. (ed.) (2012) *The SAGE Handbook of Interview Research: The Complexity of the Craft*. London: Sage.
- Hakama M, Malila N and Dillner J (2012) Randomised health services studies. *International Journal of Cancer* 131(12), 2898–902.

- Holma K and Kontinen T (2011) Realistic evaluation as an avenue to learning for development NGOs. *Evaluation* 17(2), 181–92.
- Ilmakunnas S, Junka T and Uusitalo R (eds) (2008) *Vaikuttavaa tutkimusta: miten arviointitutkimus palvelee päätöksenteon tarpeita?* A study on evaluative economic research for decision-making, in Finnish. Helsinki: Government Institute for Economic Research.
- INAHTA (International Network of Agencies for Health Technology Assessment) (2013) Homepage. URL: <http://www.inahta.org>
- INTOSAI (International Organization of Supreme Audit Institutions) (2007). *INTOSAI Working Group on Program Evaluation, Program Evaluation for SAIs: A Primer*. London: National Audit Office.
- Jordanova L (2006) *History in Practice*. London: Bloomsbury Academic.
- Keskitalo P, Määttä K and Uusiautti S (2012) Sámi education in Finland. *Early Child Development and Care* 182(3–4), 329–43.
- Kruecken, G. and Drori, G. (ed.) (2009). *World Society: The Writings of John W. Meyer*. Oxford: Oxford University Press.
- Laitinen, I. (2008) *Autenttinen arviointietiikka*. A study in the ethics of evaluation, in Finnish. Rovaniemi: University of Lapland.
- Lowndes V and Roberts M (2013) *Why Institutions Matter: The New Institutionalism in Political Science*. Houndmills: Palgrave Macmillan.

MEC (Ministry of Education and Culture) (2013a) Koulutuksen arviointikeskuksen perustaminen. On founding the Center of Education Evaluation, in Finnish. URL: http://www.minedu.fi/OPM/Koulutus/koulutuspolitiikka/vireilla_koulutus/

MEC (Ministry of Education and Culture) (2013b) OECD Pisa Survey. URL: <http://www.minedu.fi/OPM/Koulutus/artikkelit/pisa-tutkimus/?lang=en>

Meyer JW and Jepperson RL (2000) The ‘actors’ of modern society: the cultural construction of social agency. *Sociological Theory* 18(1): 100–20.

Meyer JW and Rowan B (1977) Institutional organizations: formal structure as myth and ceremony. *American Journal of Sociology* 83(2): 340–63.

Meyer RE and Hammerschmid G (2006) Changing institutional logics and executive identities: a managerial challenge to public administration in Austria. *American Behavioral Scientist* 49(7): 1000–1014.

MF (Ministry of Finance) (2003) *Parempaan tilivelvollisyyteen: valtion tilinpäätösuudistuksen periaatteet*. Report on the reform of government accounting closure, in Finnish. Helsinki: MF.

MF (Ministry of Finance) (2012) *Uudistamme yhdessä valtionhallintoa*. Brochure on the government program on effectiveness and performance, in Finnish. Helsinki: MF.

MF (Ministry of Finance) (2013a) Peruspalvelujen arviointi. The evaluation of the basic services, in Finnish. URL:

http://www.vm.fi/vm/fi/13_hallinnon_kehittaminen/03_palvelujen_laatu/

MF (Ministry of Finance) (2013b). *Valtioneuvoston selonteko eduskunnalle aluehallintouudistuksen toimeenpanosta ja toteutumisesta*. Government report to Parliament on the effects of the reform of regional administration, in Finnish. Helsinki: MF.

MFA (Ministry for Foreign Affairs) (2013) Evaluations. URL:

<http://formin.finland.fi/public/default.aspx?nodeid=15454&contentlan=2&culture=en-US>

MJ (Ministry of Justice) (2004) *Hallituksen esitysten laadintaohjeet*. Guidelines for preparing government proposals to Parliament, in Finnish. Helsinki: MJ.

OECD (2010) *Finland: Working Together to Sustain Success*. Paris: OECD.

Patton MQ (2008) *Utilization-Focused Evaluation*. 4. ed. London: Sage.

Pawson R and Tilley N (1997) *Realistic Evaluation*. London: Sage.

Peters BG (2011) *Institutional Theory in Political Science: The 'New Institutionalism'*. 3. rev. ed. London: Continuum.

PMO (Prime Minister's Office) (2011) *Poliittisen toiminnan tietopohjan parantaminen*. On improving the information basis of political decisions, in Finnish. Helsinki: PMO.

PMO (Prime Minister's Office) (2013a) PMO budget proposal for 2014, in Finnish.

URL: <http://vnk.fi/ajankohtaista/tiedotteet/tiedote/fi.jsp?oid=391631>

PMO (Prime Minister's Office) (2013b) *Hallitusohjelman tavoitteiden ja toimien toteutuma*. Mid-term evaluation of the government political program, in Finnish. Helsinki: PMO.

Pollitt C (ed.) (2013) *Context in Public Policy and Management*. Cheltenham: Edward Elgar.

Pollitt, C and Bouckaert G (2011) *Public Management Reform: A Comparative Analysis – New Public Management, Governance, and the neo-Weberian State*. 3. ed. Oxford: Oxford University Press.

Pollitt C, Hanney S, Packwood T, Rothwell S and Roberts S (1997) *Trajectories and Options: An International Perspective on the Implementation of Finnish Public Management Reforms*. Helsinki: Ministry of Finance.

Pope J, Bond A, Morrison-Saunders A and Retief I (2013) Advancing the theory and practice of impact assessment: setting the research agenda. *Environmental Impact Assessment Review* 41(1), 1–9.

Renko M (1971) *Opettajan tehokkuus oppilaiden koulusaavutusten ja persoonallisuuden kehittämisessä*. A study on teacher efficiency, in Finnish. Jyväskylä: University of Jyväskylä.

- Salminen A, Viinamäki O-P and Jokisuu J (2012) Finland. In: Verhoest K, Van Thiel S, Bouckaert G and Lægreid P (eds) *Government Agencies: Practices and Lessons from 30 Countries*. Houndmills: Palgrave Macmillan, 223–33.
- Salusjärvi M and Kontiala P (1975) *Tiekohtaisten nopeusrajoitusten kokeilu 1973: vaikutukset liikenneonnettomuuksiin*. A quasi-experimental study on speed limits, in Finnish. Otaniemi: VTT.
- Schneiberg M and Clemens ES (2006) The typical tools for the job: research strategies in institutional analysis. *Sociological Theory* 24(3), 195–227.
- Schwandt TA (2002) *Evaluation Practice Reconsidered*. New York: Peter Lang.
- Silverman D (2011) *Interpreting Qualitative Data*. 4. ed. Los Angeles, CA: Sage.
- Sintonen H (1981) *An Approach to Economic Evaluation of Actions for Health*. Helsinki: Ministry of Social Affairs and Health.
- Strang D and Macy MW (2001) In search of excellence: fads, success stories, and adaptive emulation. *American Journal of Sociology* 107(1), 147–82.
- Stufflebeam, DL and Shinkfield AJ (2007) *Evaluation Theory, Models, and Applications*. San Francisco, CA: Jossey-Bass.
- Tala J (2010) Better regulation through programs and quality standards – are new perspectives needed? *Legisprudence* 4(2), 193–212.
- Uusikylä P, Ahonen P and Takanen A (2015) *EU-aloitteiden vaikutusten ennakoarvioinnin tehostaminen*. A study on improving the ex ante assessment of

European Commission legal and policy initiatives in Finland. Helsinki: Prime Minister's Office.

Weiss CH (1970) The politicization of evaluation research. *Journal of Social Issues* 26(4), 57–68.

Zucker, LG (1977) The role of institutionalization in cultural persistence. *American Sociological Review* 42(5), 726–43.

Table 1. Evaluation vocabulary in Finnish legislation, 2014.

Words and expressions in Finnish		Closest equivalents in English		Number of acts or statutes with the word or expression	
Word or words	Expression	For words	For expressions	Current acts or statutes	Original acts or statutes
<i>arvioin*</i>	N.A.	evaluation	N.A.	399	1,980
arvioin* AND vaikut*	<i>vaikutta- vuuden arviointi, vaikutta- vuus- arviointi</i>	evaluation AND effectiveness	effectiveness evaluation, impact evaluation	315	976
arvioin* AND yhteiskunnall* AND vaikut*	<i>yhteiskunnallisen vaikutta- vuuden arviointi</i>	evaluation AND societal AND effectiveness	evaluation of societal effectiveness	42	88

Source: Finlex database, 11 December 2014. Asterisks (*) indicate truncation to catch words in all their linguistic forms. AND comprises a mere connector.

Appendix 1. Aspects of the institutionalization of evaluation in the national government of Finland.

	I Characterization	II Comments	III Estimated emphasis on rationality or legitimation	IV Estimated coupling to decision-making	V Examples of origins of institutional elements
1 Parliament	Commonly obliges government to present ex post evaluations of parliamentary legislation	N.A.	Emphasis on rationality limited by the short term covered by the evaluations	Formally tight, but varies because of Finland's multi-party rule	Implements global institutional models and scripts; e.g., MF, 2013b
- 1.1 Audit Committee	May commission studies with evaluation characteristics	N.A.	Primarily legitimation	The Audit Committee is none of the committees with foremost legal preparation roles	Created in 2007 analogously as in 2003 in Sweden
- 1.2 State Audit Office (SAO)	Carries out performance auditing and compliance auditing; since 2001 no longer affiliated with government but Parliament	Many performance audits analogous with evaluations, but SAO lacks the mandate to use the term	Mandate accentuates rationality	Independent, excellent public image, but coupling to decision-making rather loose despite re-audit activities	Networking with its global peer institutions (e.g., INTOSAI, 2007)
- 1.3. Social Insurance Institution	May carry out studies with evaluation characteristics	N.A.	Rationality rather than legitimation	Loose except for evaluations of functions under its own executive responsibility	Cooperates with its global peer institutions
2 Prime Minister's Office	N.A.	N.A.	By and large, see points 2.0.1–2.0.5	N.A.	N.A.
	2.0.1 Evaluates the implementation of the government political	N.A.	Rationality orientation limited and legitimation orientation enhanced by the four-year		Influenced by global institutional models and scripts; e.g., PMO,

	program		electoral cycle		2013b
	2.0.2 Together with MF (see 7 below) loosely coordinates evaluation activities in government	N.A.	Accents on legitimation rather than rationality	Coupling to decision-making generally rather loose	PMO, 2011; MF, 2012
	2.0.3 Outsourced ex ante assessments of European Commission legal and other initiatives, since 2015	Both 2.0.3 and 2.0.4 tap a new funding medium of the PMO introduced in 2014, growing to 12.5 million euros p.a. by 2016	Accents on legitimation rather than rationality	Intended coupling to decision-making tight rather than loose	Uusikylä et al., 2015
	2.0.4 Other possible outsourced assessments and evaluations		N.A.	N.A.	PMO, 2013a
	2.0.5 Evaluation components in studies funded by Council of Strategic Research according to principles defined by government	Another new funding medium; will grow to 63 million euros p.a. by 2016; the first projects start in 2015	N.A.	N.A.	PMO, 2013a
3 Ministry for Foreign Affairs	Evaluation of Finland's development co-operation, ever since the 1970s	N.A.	Legitimation (intangibles), rationality (program and project management, funding)	Looser (intangibles, legitimation) or tighter (tangibles, rationality)	Harmonized with global and EU guidelines; MFA, 2013
4 Ministry of Justice	Maintains guidelines for the ex ante evaluation of the impacts of all proposed new	The guidelines only indicate the impact types to consider	Legitimation (intangibles), in principle also rationality	Made looser by the absence of the compulsory operationalization of	Influenced by global models and scripts of legal impact evaluation; see MJ, 2004

	legislation		(tangibles)	the impacts		
- 4.1. Institute for Legal Policy Research	Carried out studies with evaluative characteristics	Formally a unit of the University of Helsinki since 2015, but retains previous functions	Legitimation (intangibles), rationality (processes of legal preparation and implementation)	Made ever looser by new status within a university faculty and strengthening emphasis upon issues of criminology	Networks with its global peer institutions	
5 Ministry of the Interior	N.A.	N.A.	N.A.	N.A.	N.A.	
6 Ministry of Defense	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
7 Ministry of Finance	N.A.	N.A.	By and large, see items 7.0.1–7.0.3 and 7.1		N.A.	
	7.0.1 Regulates reporting on the results of evaluations carried out by the ministries, agencies and offices	N.A.	Rationality and legitimation	Made looser by information overload and the high cognitive demands of effectiveness evaluation	Influenced by global models and scripts of evaluation and extended accounting; BA, 2013; BS, 2013	
	7.0.2 Coordinates the evaluation of the basic public services	N.A.	Legitimation (intangibles), performance (tangibles)	Loose, as the service providers mostly comprise self-governing municipal organizations	MF, 2013a	
	7.0.3 Evaluation of public administration reform policies	N.A.	Both legitimation and performance	Loose rather than tight	Active in the 1990s (e.g., Pollitt et al., 1997), later abated	
- 7.1 Government Institute for	Characterizes some of its work as policy	E.g., economic evaluations of	Entrenched provider of studies with	Made looser because of staff instead of line	Ilmakunnas et al., 2008	

Economic Research	evaluation	social programs	evaluation characteristics	organization status, but lately turned visibly tighter	
8 Ministry of Education and Culture	N.A.	N.A.	By and large, see points 8.1 and 8.2	N.A.	
- 8.1 Center of Education Evaluation	In 2014, assumed responsibilities of evaluation types 8.1-8.3 from their previous provider				MEC, 2013a
- 8.1.1 Evaluation of compulsory education	Studies of the educational achievement of students finishing their compulsory education	Part of the OECD-PISA studies	Rationality, but spillovers to legitimation from Finland's excellent results	Coupling made looser by prohibition to publish the results by schools	The PISA studies are an OECD project; see MEC, 2013b
- 8.1.2 Evaluation of vocational and related education	Evaluation of vocational education and adult education	N.A.	Varies	Loose rather than tight	N.A.
- 8.1.3 Higher education evaluation	Audits of quality assurance systems of higher education, evaluations by the Center of Education Evaluation or the universities themselves with Center assistance	Legislation obliges the polytechnics and the universities to be subject to the evaluations	Legitimation rather than rationality	Generally loose	Influenced by global institutional models and scripts; ENQA, 2013
- 8.2 The Academy of Finland	Carries out evaluations of research policies and research fields	Comprises scientific boards and an administrative office	Rationality rather than legitimation	Coupling varies, made looser by staff as opposed to line status of the AF boards and the AF office	Finland's science policies have been strongly influenced by global and EU models and scripts

9 Ministry of Agriculture and Forestry	See the fourth column	Has important research institutes within its sector of administration	The institutes are entrenched providers of studies with evaluation characteristics	The staff as opposed to line status of the institutions does not prevent their reasonably tight coupling with decision-making	N.A.
10 Ministry of Transport and Communications	N.A.	N.A.	Rationality rather than legitimation	Varies	N.A.
11 Ministry of Employment and the Economy	N.A.	N.A.	Too diffuse to indicate in detail; however, let us mention that the ministry carries out regulatory impact assessment (RIA) using the common standard cost model to assess regulatory burdens		RIA derives from global institutional models and scripts
- 11.1 Technological Research Institute	May carry out evaluation studies on commission	N.A.	Entrenched provider of studies with evaluation characteristics	Looser because of staff instead of line organization status	Following international examples, will turn into a government-owned company in 2015
- 11.2 Funding Agency for Technology and Innovation	Has the programs and projects that its funds systematically evaluated	N.A.	There is too much variance in these aspects to allow for estimation		Evaluation much influenced by EU institutional models and scripts
12 Ministry of Social Affairs and Health	Among its other duties, maintains within its organization the Finnish Office for Health Care Technology Assessment (FINOHTA)	Promotes impact assessment related to health, welfare, gender, social and human aspects	More rationality orientation in FINOHTA, more legitimation orientation in the other assessment activities	Tighter where rationality receives emphasis, looser where legitimation is important	Health technology assessment in Finland derives from global models and scripts (INAHTA, 2013)

- 12.1 National Institute of Health and Welfare	N.A.	Carries out evaluations of health care and social welfare, including in its FinSOC unit for social work evaluation	Entrenched provider of studies with evaluation characteristics	Loose rather than tight; generally tighter in health than in social welfare	N.A.
- 12.2 Finnish Institute of Occupational Health	Carries out academic and applied evaluation studies	N.A.	Entrenched provider of studies with evaluation characteristics	Made looser by the corporatist influence of the employees and employers in occupational health	Connected with its global counterpart organizations
13 Ministry of the Environment	N.A.	A foremost actor in statutory environmental impact assessment (EIA)	Rationality (in principle), legitimation (where evaluation is more of a formality)	Looser where vested interests are strong	EIA is a global innovation (Pope et al., 2013) diffused in Finland
- 13.1 Finnish Environmental Institute	Carries out academic and applied evaluation studies	N.A.	Entrenched provider of studies with evaluation characteristics	Looser because of staff instead of line status	N.A.

Explanation: Subordinate bodies of functions are indicated by a dash below the name of their main body.

Appendix 2. Thematic prompts in interviews of Finnish evaluation experts.

(1) What is the basic institutionalization of evaluation in Finland like?

(1.1) How is evaluation understood in Finland?

(1.1.1) As such

(1.1.2) In relation to other, resembling notions and practices

(1.1.3) In relation to different notions and practices

(1.2) Is *arviointi* in Finnish unequivocally equal with the Swedish *utvärdering* and the English ‘evaluation’? How do you conceive of the relationship of ‘evaluation’ and *arviointi* to ‘assessment’ and ‘inspection’? Does *tuloksellisuustarkastus* (‘performance auditing’) refer to *arviointi*?

(1.3) Does the Finnish language adequately demarcate evaluation from what it is not? Do you have observations on overexpansion or excessive contraction?

(1.4) Can you outline the identity of evaluation in Finland in distinction from what it is not and, in addition, outline systematic development trends in this identity?

(2) How do you conceive of the ‘agency’ (‘agenssi’, *toimijuus*) in evaluation in Finland?

(2.1) How do you conceive of the agency of Finland’s organizations and institutions of evaluation for themselves and in themselves, such as their independence on the one hand, and on the other their dependence on the organizations commissioning evaluations?

(2.2) How do you conceive of the agency of those who carry out or commission evaluations? Such as:

(2.2.1) procurement competency

(2.2.2) capability and willingness to maintain the independence of evaluation

(2.2.3) evaluation ethics in general

(2.3) How do you conceive of the application of global and other principles and standards of evaluation in Finland? Such as:

(2.3.1) the critical and competent application of global principles and standards

(2.3.2) the professionalization of evaluation and the consequences of professionalization

(2.3.3) the general ethical quality of evaluation in Finland

(3) Driven by its theoretical background, this study pays attention to relationships between rationality (e.g., efficiency, performance, and effectiveness) and legitimation, including possible complementary relationships between these two aspects. How do you conceive of this issue?

(3.1) Please mention cases in which the adoption and application of global institutional models to set up institutions, organizations and practices of evaluation in Finland have been influenced by:

(3.1.1) aims to advance rationality

(3.1.2) aims to advance legitimation

(3.1.3) both aims

(3.2) Please mention cases, in which the launching and implementation of evaluation programs or projects has been influenced by:

(3.2.1) aims to advance rationality

(3.2.2) aims to advance legitimation

(3.2.3) both aims

(3.3) Please mention cases in which evaluation and their results (findings, valuations, recommendations, etc.) have been coupled with institutions and practices

(3.3.1) in a tight way (e.g., by means of legal sanctions or conditional future funding)

(3.3.2) in a loose way

(3.3.3) hardly or not at all

(4) Can you characterize periods or moments – consider times since 1987 – that evaluation has been the object of intense institutional change in Finland? Consider only changes in institutional models, practices, principles and techniques of evaluation, and the application of the evaluation results. The

changes I want to trace are those that have made earlier elements unpopular and may have led to their replacement by more popular successor elements.

(4.1) Do you find that the period 1987–91 or some parts of it can be characterized in the way indicated above? In these years, Harri Holkeri was the Prime Minister and his party, the Conservatives, was a government coalition party for the first time since 1966. You might, for instance, pinpoint:

(4.1.1) The effects of the introduction of result-oriented steering and result-oriented budgeting

(4.1.2) The onset of pro-market thinking

(4.2) Do you find that the period 1991–7 or some of its parts can be characterized in the way indicated above? By way of reminder, here are some contemporary events: the dissolution of the Soviet Union (1991), the deepest trough of Finland's economic crisis (1992–4), Finland's application for EU membership (1992), and the start of EU membership (1995). You might, for instance, pinpoint:

(4.2.1) The lagged effects of the 1987–91 reforms upon evaluation during 1991–7

(4.2.2) The effects of the grave economic crisis

(4.2.3) Effects of Finland's EU membership upon evaluation